DOCUMENT RESUME

ED 424 804 HE 031 599

AUTHOR Borden, Victor M. H.; Dalphin, John F.
TITLE Simulating the Effect of Student Profile Changes on Retention and Graduation Rates: A Markov Chain Analysis. AIR 1998 Annual Forum Paper.
PUB DATE 1998-05-19
NOTE 21p.; Paper presented at the Annual Forum of the Association for Institutional Research (38th, Minneapolis, MN, May 17-20, 1998).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Persistence; College Students; Credits; Enrollment; *Grade Point Average; *Graduation; Higher Education; *Markov Processes; Models; Predictor Variables; *School Holding Power; Simulation; Student Characteristics; Universities
IDENTIFIERS *AIR Forum

ABSTRACT

This study used Markov chain matrices to simulate the effect of varying degrees of change in student characteristics on retention and graduation rates. Data were applied to a 1-year enrollment transition matrix that tracks how students of each class level progress into the same or higher class levels, to a completed degree, or to non-returning status. The results indicated that there are large, initial differences in graduation rates according to grade-point average (GPA) and course credit-load differences among students. Despite the strong association between grades and retention, this analysis indicated that even a 25 percent upward shift in grade distribution increased the total graduation rate by only 2.4 percent. A shift in 25 percent of students toward higher credit-load categories resulted in a total increase in the graduation rate of less than 1 percent. The simulation demonstrates that even significant changes in the distribution of students by GPA and course credit-load do not yield significant changes in graduation rates. (MDM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 424 804

Running Head: SIMULATING STUDENT PROFILE CHANGES

Simulating the Effect of Student Profile Changes on Retention and Graduation Rates: A Markov Chain Analysis

Victor M. H. Borden

Director and Assistant Professor

Indiana University-Purdue University Indianapolis

620 Union Drive, UN G003

Indianapolis, IN 46202-5167

Phone: (317) 274-8213

John F. Dalphin

Assistant to the Chancellor and Professor

Indiana University East

2325 Chester Blvd

Richmond, IN 47374

Phone: (765) 973-8530

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

AIR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Paper presented at the 38th Annual Forum of the Association for Institutional Research

Minneapolis, MN - May 19, 1998.

IR 031 599

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Eighth Annual Forum of the Association for Institutional Research held in Minneapolis, Minnesota, May 17-20, 1998.
This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of AIR Forum Papers.

Dolores Vura
Editor
AIR Forum Publications

Simulating the Effect of Student Profile Changes on Retention and Graduation Rates: A Markov Chain Analysis

Abstract

Markov chain matrices can be applied to student enrollment transitions in a variety of ways to model and forecast student flow. In this study, Markov processes are used to simulate the effect of varying degrees of change in student characteristics on retention and graduation rates. More specifically, changes in student academic performance and course credit load are considered insofar as they are linked closely to traditional persistence and graduation rates. Results show that even though there is a strong association between grade performance and persistence, it takes very large changes in levels of student performance to impact retention and graduation rates modestly.

Simulating the Effect of Student Profile Changes on Retention and Graduation Rates: A Markov Chain Analysis

Undergraduate student retention and graduation rates have taken on increasing importance as barometers of institutional effectiveness for U.S. colleges and universities. Federal regulations such as the Student Right-to-Know Act (SRKA), NCAA reporting requirements, state-level funding initiatives, and commercial college rankings such as those of *U.S. News & World Report* have contributed, for better or worse, to the preeminence of these measures.

Like most measures related to higher education, retention and graduation rates were developed to reflect the traditional college experience: full-time attendance at a residential college among college-prepared high school graduates with few concurrent obligations. Nationwide declines in retention and graduation rates may well reflect the inadequacy of the traditional model more than changes in institutional effectiveness.

In addition to the decreasing relevance for many postsecondary institutions, graduation rates do not lend themselves readily to analyzing the effectiveness of programs intended to increase student persistence to graduation. That is, a six-year rate requires a time lag that renders it

difficult to judge current program effectiveness. At the same time, six years may be an insufficient time frame to reflect completion for students who attend at less than a full-time rate through their academic career. Markov chain models have been demonstrated to provide an effective means for measuring student progress that circumvents this time-lag concern (Dalphin, 1997; Donhardt, 1995; Tukey, 1991).

This paper presents a Markov chain analysis of student progress that employs a one-year transition matrix that tracks how students of each class level progress into the same or higher class levels, to a completed degree, or to a non-returning status. Matrix calculations are performed to determine a variety of progression rates. The resulting rates include the average time it will take to reach one state (e.g., a degree) from another (e.g., freshmen), as well as the percentage of individuals in one state (e.g., freshmen) who will eventually reach a "terminal" state (e.g., a degree).

Building upon other research conducted by the authors (Dalphin, 1997; Dalphin and Borden, in preparation), this study focuses on the use of discrete Markov chain processes to simulate the effect of changes in student body profile on graduation rates. Since students' course credit-load (e.g., full- versus part-time status) and grade performance are known to be strong correlates of retention (Pascarella & Terenzini,

1991), this study considers how changes in student profile along these two dimensions affects graduation rates.

A Markov Chain Analysis of Student Transitions

Like all Markov chain analyses, the current model of student flow employs a transition matrix that summarizes the probabilities of students moving from one 'state' to another between two points in time. Before employing the usual Markov matrix calculations to determine various rates of transition, the authors devised a strategy for incorporating stopout enrollments, that is, those students who fail to re-enroll during the follow-up period but will later re-enroll and resume their studies. Markov matrix calculations were then applied to the adjusted transition matrix, providing the basis for simulating the impact of changes in student grade performance on graduation rates.

The Transition Matrix

The authors previously developed a model for deriving graduation rates using Markov-chain process as applied to a one-year enrollment transition matrix (Dalphin, 1997; Dalphin & Borden, in preparation)[1]. In its simplest form, students are arrayed into the matrix according to their class level

(freshman, sophomore, etc.) during the base year (as indicated in the rows of the matrix) and by their class level or among several non-enrolled categories during the follow-up year (as indicated by the column position). Table 1 illustrates this simple form of the transition matrix, showing the number of students who fall into each cell of the matrix (top portion) as well as the transition rates (bottom portion).

Table 1. A Simple Transition Matrix for A Markov-Chain Analysis of Student Flow

Base Semester	Follow-up Semester						Total
	Freshman	Sophomore	Junior	Senior	Not Enrolled	Graduated	
Number							
Freshman	1410	1257	27	1	2430	0	5125
Sophomore	0	886	1110	62	1066	0	3124
Junior	0	0	668	1462	761	30	2921
Senior	0	0	0	1990	1063	1838	4891
Percent							
Freshman	28%	25%	1%	0%	47%	0%	100%
Sophomore	0%	28%	36%	2%	34%	0%	100%
Junior	0%	0%	23%	50%	26%	1%	100%
Senior	0%	0%	0%	41%	22%	38%	100%
	Q Matrix				**R Matrix**		

The transition rates, shown as percentages in the bottom panel of Table 1, are the primary input into the Markov calculations. Specifically, the rates of transition from one non-absorption state into another (e.g., freshman to sophomore) comprise the *Q* matrix and transitions from non-absorption states into absorption (i.e., terminal) states comprise the *R* matrix. Using matrix calculations, the Markov chain process takes the transition rates through continuing iterations, *ad infinitum*.

That is, the students who continue in non-absorption states are processed through the matrix using the same rates of transition, until asymptotically all students reach a final absorption state: either graduated or permanently not enrolled. One Markov matrix formula can be used to determine the average time it takes to reach an absorption state from each of the initial non-absorption state. Of more direct interest in this study is the matrix formula used to determine the percentage of students from each initial non-absorption state that will eventually reach the various absorption states, and especially the state of "graduated." The formula for deriving this matrix is,

$$\mathbf{B} = (\mathbf{I} - \mathbf{Q})^{-1} * \mathbf{R}$$

where **I** is an identity matrix equal in rank to the **Q** matrix.

In preparation for the simulation study, the transition matrix was expanded to incorporate the credit-load and grade performance categories. In addition, the non-enrolled category was expanded to accommodate various sub-states that were relevant to the authors' institutions. These included the ability to track students across various campuses of the university, as well as enrollment as a non-degree student or attainment of an associate's degree, which may or may not be the 'terminal' degree for a given program.

Table 2 summarizes the final set of variables and their categorical values, used to array students in the base (rows)

and follow-up (columns) semesters of the expanded transition matrix. For purposes of simplicity and communication, credit-load and grade performance each were considered in three categories. The matrix employs identical class level, credit-load, and grade performance categories for both the base and follow-up semester status among enrolled students. The follow-up semester additionally includes four 'non-degree-enrolled' categories as well as the 'terminal completion' category of receiving the bachelor's degree.

Table 2. Base Year (Row) and Follow-up Year (Column) Status Categories of Expanded Transition Matrix

Variable	*Values*
Year 1 and Year 2 Enrolled Status Categories (Row and Column)	
Class Level	Freshman, Sophomore, Junior, Senior
Semester Credit-load	1 – 5 credits; 6 – 11 credits; 12 or more credits
Grade Performance (GPA)	< 2.00; 2.00 – 2.99; 3.00 or higher (4-point scale)
Year 2 only Non-Degree-Enrolled Status Categories (Column only)	
Not enrolled in degree-seeking program nor received bachelor's degree	Enrolled at other system campus; enrolled as non-degree student; not enrolled anywhere that can be tracked; received associate degree
Received bachelor's degree	(terminal graduation status for undergraduate degree programs)

Figure 1 provides an abbreviated representation of the resulting transition matrix when all categories are considered simultaneously.

BEST COPY AVAILABLE

Figure 1. Expanded Transition Matrix to Support Simulation (Abbreviated)

Initial Fall Status		Class-->	Enrolled in Undergraduate Degree-Seeking Program the Next Fall Semester									Not Enrolled in Ugrad Degree Program					
		Credits-->	Freshman								Sophomore - Senior	Enrolled Other Campus	Enrolled Non-Degree Program	Not Enrolled Anywhere	Received Associate Degree	Received Bacc. Degree	Row Total
			0 - 5			6 - 11			12 +								
Class	Credits	GPA	<2.0	2.0-2.99	3.0 +	<2.0	2.0-2.99	3.0 +	<2.0	2.0-2.99	3.0 +						
Freshman	0 - 5	<2.0	14	9	2	9	2	2	9	0	1	0	0	167	0	0	223
		2.0-2.99	3	30	8	2	10	2	1	4	1	1	0	97	0	0	181
		3.0 +	0	11	59	1	10	35	1	3	4	2	2	171	0	0	331
	6 - 11	<2.0	12	2	0	43	21	0	22	8	1	5	0	449	0	0	618
		2.0-2.99	7	14	2	26	49	6	15	28	2	8	0	179	0	0	483
		3.0 +	0	5	23	8	37	82	3	20	25	14	2	133	0	0	489
	12 +	<2.0	8	2	0	50	8	2	135	42	1	10	0	672	0	0	1023
		2.0-2.99	2	9	0	17	36	3	88	134	5	70	0	271	0	0	1015
		3.0 +	1	2	5	0	10	11	18	75	54	49	0	128	0	0	762
Junior - Sophomore																	
Senior	0 - 5	<2.0	0	0	0	0	0	0	0	0	0	0	1	18	0	5	39
		2.0-2.99	0	0	0	0	0	0	0	0	0	3	5	194	6	87	487
		3.0 +	0	0	0	0	0	0	0	0	0	5	12	161	1	82	421
	6 - 11	<2.0	0	0	0	0	0	0	0	0	0	0	0	44	0	1	72
		2.0-2.99	0	0	0	0	0	0	0	0	0	5	3	158	9	215	761
		3.0 +	0	0	0	0	0	0	0	0	0	5	22	84	9	232	642
	12 +	<2.0	0	0	0	0	0	0	0	0	0	0	0	32	0	4	62
		2.0-2.99	0	0	0	0	0	0	0	0	0	11	5	122	13	429	1037
		3.0 +	0	0	0	0	0	0	0	0	0	5	9	111	10	783	1388

Accounting for Stop-Outs

One limitation of using a one-year transition matrix to reflect student flow is the inability to capture the phenomenon commonly called stopouts: those students who re-enroll in a college or university after a period of absence. To account for stopouts in this model, re-entry rates were calculated from among students from an historical cohort. Specifically, re-entry rates were developed using students who enrolled as undergraduate degree-seeking students during the fall 1987 semester but did not enroll nor receive a baccalaureate degree by fall 1988. These students were arrayed according to the their fall 1987 enrollment status (by class level, credit-load and GPA) and fall 1988 non-returning status (enrolled at other campus, enrolled non-degree, not enrolled anywhere and received associate degree). The re-entry rate was then determined for

BEST COPY AVAILABLE

each cell of this matrix according to the percentage of students who subsequently re-enrolled after the fall 1988 semester, up through the most recently available semester at the time of the study (spring 1997). Thus the re-entry rate is based on a ten-year time period. Figure 2 illustrates the re-entry rates determined from this historical cohort.

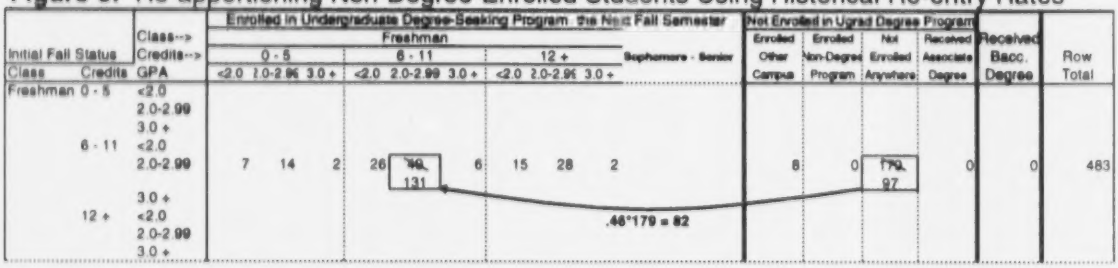
Figure 2. Re-entry Rates For Historical Cohort (Abbreviated)

Initial Fall Status			Percent of Corresponding Students from Fall 1987			
Class	Credits	GPA	Enrolled Other Campus	Enrolled Non-Degree Program	Not Enrolled Anywhere	Received Associate Degree
Freshman	0 - 5	<2.0	75%	0%	26%	0%
		2.0-2.99	50%	0%	41%	0%
		3.0 +	0%	0%	32%	0%
	6 - 11	<2.0	60%	0%	31%	0%
		2.0-2.99	46%	100%	46%	0%
		3.0 +	55%	100%	41%	0%
	12 +	<2.0	45%	0%	42%	0%
		2.0-2.99	55%	0%	39%	0%
		3.0 +	46%	0%	36%	0%
Sophomore - Junior						
Senior	0 - 5	<2.0	0%	0%	35%	0%
		2.0-2.99	0%	25%	43%	33%
		3.0 +	17%	0%	33%	0%
	6 - 11	<2.0	0%	0%	36%	0%
		2.0-2.99	100%	20%	41%	67%
		3.0 +	17%	0%	40%	100%
	12 +	<2.0	0%	100%	39%	0%
		2.0-2.99	60%	0%	38%	67%
		3.0 +	50%	0%	38%	0%

The re-entry rates were used to apportion students within each of the non-enrolled categories in the returning year (columns) of the recent transition matrix among those who would be forever lost and those who would subsequently return. The 'forever lost' portion remained in the non-enrolled category as

a terminal state and the portion that was estimated to return was placed back into the column that corresponded to the row in which they started. For example, of all the fall 1987 freshman taking 6-11 credits and achieving a first semester GPA between 2.00-2.99 GPA who did not re-enroll anywhere in fall 1988, 46 percent subsequently re-enrolled at a later date as degree-seeking undergraduates (emboldened cell of Figure 2) and 54 percent never returned. Therefore, 46 percent of the students falling in the initial cell corresponding to the same row (freshman, 6 - 11 credit, 2.00 - 2.99 GPA) and column (not enrolled anywhere) of the current matrix were moved into the column representing their base year category (freshman, 6 - 11 credit, 2.00 - 2.99 GPA). That is, it was assumed that students would return into the same class level, credit-load and GPA group as when they left. Figure 3 three illustrates this process.

Figure 3. Re-apportioning Non-Degree-Enrolled Students Using Historical Re-entry Rates



BEST COPY AVAILABLE

Calculating Graduation Rates from the Transition Matrix

The resulting one-year transition matrix is converted into a rate matrix by representing each cell as a percentage of the row total. These rates then represent the percentage of students who move from each initial state, represented by the row position, to each follow-up state, represented by the column position, one year later. The Markov matrix calculations described above were then applied to the transition rates of the revised matrix to generate the Markov-chain graduation rates for all students in the base year cohort.

Table 3 shows the resulting Markov graduation rates for freshman in the base semester according to their first semester credit load and grade-point average. Multiplying these rates by the original number of students in each group provides estimates of how many from each group would reach this terminal state. These numbers can then be used to calculate graduation rates for various combinations of the original cohort as shown in the bottom portion of Table 3.

To compare the resulting graduation rates with more traditional measures, the authors used the same technique on an earlier cohort for which traditional measures were available. Among first-time full-time freshman, the traditional six-year graduation rates for the 1987 cohort was 25.3%, which compared quite closely to the corresponding Markov graduation rate for

all full-time freshmen (not just first-time) of 29.7%. Comparing the rates for part-time students, the traditional six-year rate for the 1989 entering cohort was 8.4% compared to the Markov rate of 15.4%. Since the Markov process covers a theoretically indefinite time period. It is not surprising that the Markov rates are slightly higher than the six-year rates, especially among part-time students. For the 1987 cohort, the corresponding nine-year graduation rates were 30.1% for full-time freshmen, and 13.2% for part-time freshmen, that is, even closer to the Markov-modeled rates.

Table 3. Estimating Graduation Rates from the Markov-Chain Analysis

Initial Fall Status			Received Bacc. Degree	Original N	Projected Number of Graduates
Class	Credits	GPA			
Freshman	0 - 5	<2.0	0.0349	223	7.8
		2.0-2.99	0.1152	181	20.9
		3.0 +	0.1271	331	42.1
	6 - 11	<2.0	0.0586	618	36.2
		2.0-2.99	0.2036	483	98.3
		3.0 +	0.2628	489	128.5
	12 +	<2.0	0.0785	1023	80.3
		2.0-2.99	0.2570	1015	260.9
		3.0 +	0.4147	762	316.0
Estimated Graduation Rates					
		All Freshmen	19.3%		
		Part-Time	14.4%		
		Full-Time	23.5%		
		GPA>3.00	30.8%		
		GPA 20	22.6%		
		GPA<2.00	6.7%		

Simulating the Impact of Changes in Student Body Profile

Once the final Markov transition matrix is calculated, various simulations can be conducted by altering the distribution of students in the base year rows of the matrix. The results of two specific models will be illustrated here, focusing on the transition of freshmen in the base year to the terminal state of receiving a bachelor's degree, that is, the Markov-modeled graduation rate.

In the first simulation, a proportion of students is re-distributed from the lower grade performance levels to the higher ones. The shifts were accomplished by moving a fixed percentage of students in each of the lower GPA categories (<2.00; and 2.00 - 2.99) up to the next category. These shifts were conducted in increments of five percent. Table 4 shows the results of this simulation. The first row of the table shows the baseline rates from Table 3 for all freshmen as well as among full- and part-time freshmen. The second row shows the resulting rates after moving five percent of the 'below 2.00' group up to the 2.00 - 2.99 category and five percent of the 2.00 - 2.99 category moving up to the 3.00 and above category. These shifts were implemented in parallel across the different course load categories. Subsequent rows show the effect of additional shifts of five-percent increments, up to the point

where one-quarter of each grade performance group is shifted up to the next grade performance category.

Table 4. Impact of Grade Distribution Shifts on Markov-Chain Derived Graduation Rates for Freshmen

Percent Shift in Grade Distribution	Total	Credit Load	
		Full-Time	Part-Time
None	19.3%	23.5%	14.4%
5%	19.8%	24.1%	14.7%
10%	20.3%	24.7%	15.0%
15%	20.7%	25.3%	15.2%
20%	21.2%	25.9%	15.5%
25%	21.7%	26.5%	15.8%

Despite the strong association between grades and retention, this analysis shows that even a 25 percent upward shift in grade distribution increases the Total graduation rate by 2.4 percentage points. It also shows that the impact of grade shifts is greater among full-time students (3.0%) than among part-time students (1.4%). That is, grades are more strongly correlated with retention rates among full-time students than among part-time students.

Table 5 shows the results of corresponding shifts in distribution of students from the lower course credit-load categories to the higher ones. Again, the table starts with the baseline rates from Table 3 and then shows the impact of shifts of five percent increments of students from the lower credit-load categories to the next higher ones (implemented in parallel across the three grade performance categories).

Table 5. Impact of Credit Load Distribution Shifts on Markov-Chain Derived Graduation Rates for Freshmen

Percent Shift in Credit Load Distribution	Total	GPA		
		3.00 +	2.00-2.99	< 2.00
None	19.3%	30.8%	22.6%	6.7%
5%	19.5%	31.1%	22.8%	6.7%
10%	19.7%	31.5%	22.9%	6.8%
15%	19.9%	31.9%	23.0%	6.8%
20%	20.0%	32.3%	23.1%	6.9%
25%	20.2%	32.6%	23.3%	6.9%

The impact of shifts in student distribution across course credit-load categories is even less notable than for the grade performance categories. Each shift of five percent in the student body results in about two-tenths of a percentage point increase in the graduation rate and shifting one-quarter of the students toward the higher credit-load categories results in a total increase of less than one percentage point. The credit-load shift impact is greatest among students with an average GPA of 3.00 or higher, but even here the impact is only about four-tenths of a percentage point increase per five percent shift in students.

Implications

The Markov-chain analysis of student transitions yields graduation rates that offer three distinct advantages over traditional rates. First, the method employs progress rates among recently enrolled students. Second, one can include all students from within the student population--new, returning, and

transfers, freshmen through seniors and so on--and thus better represent the entire student body. Finally, the method can accommodate and evaluate any grouping characteristics one believes are important in distinguishing rates of progress, and can then model the impact of changes in these characteristics on student progress.

The results of this Markov analysis shows that there are large initial differences in graduation rates according to grade performance and course credit-load differences among students. However, the simulation shows that even significant changes in the distribution of students by grade performance and course credit-load do not yield significant changes in graduation rates. These results do not bode well for those who seek to increase institutional graduation rates appreciably by making slight or even moderate changes in the selection of students according to their likelihood of enrolling as full-time students, or according to their past academic record. This is especially true with respect to increased academic selectivity, since past academic record is a less than perfect predictor of subsequent academic performance.

The current analysis does not necessarily suggest that retention programs that seek to improve student grade performance or increase the likelihood that students will enroll full-time can not be effective. On the contrary, it suggests

that factors beyond performance and course load are important. Many retention programs address more directly factors such as psychological and social commitment and feelings of identity and community, which have been shown to be more closely related to retention to graduation than simple grade or credit-load characteristics (Kuh, et. al., 1991; Pascarella & Terenzini, 1991). If anything, the current study suggests that such retention efforts should be targeted broadly to the student body of an institution, and not just to groups of at-risk students.

References

Bradley, I. & Meek, R. L. (1986). The mathematics of absorbing Markov chains. Matrices and Society. Princeton, NJ: Princeton University Press.

Dalphin, J. F. (1997). Discrete Markov matrices applied to student flow models. Paper presented at the 11th annual conference of the Indiana Association for Institutional Research. Nashville, Indiana, March 29, 1997.

Donhardt, G. L. (1995). Tracking student enrollments using the Markov chain, comprehensive tool for enrollment management. Journal of College Student Development, 36(5), pp. 457-462.

Kuh, G. D., Schuh, J. H., Whitt, E. J., & Associates (1991). Involving Colleges. San Francisco: Jossey-Bass.

Gillespie, D. T. (1992). Markov Processes: An Introduction for Physical Scientists. Boston: Academic Press.

Pascarella, E. T., & Terenzini, P. T. (1991). How College Affects Students: Findings and Insights From Twenty Years of Research. San Francisco: Jossey-Bass.

Tukey, D. D. (1991). Models for student retention and migration. Journal of the Freshman Year Experience, 3(2), 61-74.

Footnotes

[1] For a more general treatment on Markov processes, the reader should consult Bradley & Creek (1986) or Gillespie (1992).